

TTGTAACAGA AAATTTAAAT ATACTCCACT CAAGGGAATT CTGTACTTTG CCCTTTTGGT -99
 AAAGTCTCAT TTACATTTCT AAACCTTTCT TAAGAAAATC GAATTTTCCTT TGATCTCTCT -39
 1 TCTGAATTGC AGAAATCAGA TAAAAACTAC TTGGTGAA ATG ACT TCT TGT CAC ATT 18
 7 A E E H I Q K V A I F G G T H G
 GCT GAA GAA CAT ATA CAA AAG GTT GCT ATC TTT GGA GGA ACC CAT GGG 66
 23 N E L T G V F L V K H W L E N G
 AAT GAG CTA ACC GGA GTA TTT CTG GTT AAG CAT TGG CTA GAG AAT GGC 114
 39 A E I Q R T G L E V K P F I T N
 GCT GAG ATT CAG AGA ACA GGG CTG GAG GTA AAA CCA TTT ATT ACT AAC 162
 55 P R A V K K C T R Y I D C D L N
 CCC AGA GCA GTG AAG AAG TGT ACC AGA TAT ATT GAC TGT GAC CTG AAT 210
 71 R I F D L E N L G K K M S E D L
 CGC ATT TTT GAC CTT GAA AAT CTT GGC AAA AAA ATG TCA GAA GAT TTG 258
 87 P Y E V R R A Q E I N H L F G P
 CCA TAT GAA GTG AGA AGG GCT CAA GAA ATA AAT CAT TTA TTT GGT CCA 306
 103 K D S E D S Y D I I F D L H N T
 AAA GAC AGT GAA GAT TCC TAT GAC ATT ATT TTT GAC CTT CAC AAC ACC 354
 119 T S N M G C T L I L E D S R N N
 ACC TCT AAC ATG GGG TGC ACT CTT ATT CTT GAG GAT TCC AGG AAT AAC 402
 135 F L I Q M F H Y I K T S L A P L
 TTT TTA ATT CAG ATG TTT CAT TAC ATT AAG ACT TCT CTG GCT CCA CTA 450
 151 P C Y V Y L I E H P S L K Y A T
 CCC TGC TAC GTT TAT CTG ATT GAG CAT CCT TCC CTC AAA TAT GCG ACC 498
 167 T R S I A K Y P V G I E V G P Q
 ACT CGT TCC ATA GCC AAG TAT CCT GTG GGT ATA GAA GTT GGT CCT CAG 546
 183 P Q G V L R A D I L D Q M R K M
 CCT CAA GGG GTT CTG AGA GCT GAT ATC TTG GAT CAA ATG AGA AAA ATG 594
 199 I K H A L D F I H H F N E G K E
 ATT AAA CAT GCT CTT GAT TTT ATA CAT CAT TTC AAT GAA GGA AAA GAA 642
 215 F P P C A I E V Y K I I E K V D
 TTT CCT CCC TGC GCC ATT GAG GTC TAT AAA ATT ATA GAG AAA GTT GAT 690
 231 Y P R D E N G E I A A I I H P N
 TAC CCC CGG GAT GAA AAT GGA GAA ATT GCT GCT ATC ATC CAT CCT AAT 738
 247 L Q D Q D W K P L H P G D P M F
 CTG CAG GAT CAA GAC TGG AAA CCA CTG CAT CCT GGG GAT CCC ATG TTT 786
 263 L T L D G K T I P L G G D C T V
 TTA ACT CTT GAT GGG AAG ACG ATC CCA CTG GGC GGA GAC TGT ACC GTG 834
 279 Y P V F V N E A A Y Y E K K E A
 TAC CCC GTG TTT GTG AAT GAG GCC GCA TAT TAC GAA AAG AAA GAA GCT 882
 295 F A K T T K L T L N A K S I R C
 TTT GCA AAG ACA ACT AAA CTA ACG CTC AAT GCA AAA AGT ATT CGC TGC 930
 311 C L H .
 TGT TTA CAT TAG AA ATCACTTCCA GCTTACATCT TACACGGTGT CTTACAAATT 984
 CTGCTAGTCT GTAAGCTCCT TAAGAGTAGG GTTGTGCCTT ATTCAACTGC ATACATAGCT 1044
 CCTAGCACAG TGCCTTATTC GGTAGGCATC TAAGCAAATT TCTTAAATTA ATTAATATAT 1104
 CTTTAAAGAT ATCATATTTT ATGTATGTAG CTTATTCAAA GAAGTGTTTC CTATTTCTAT 1164
 ATAGTTTATT ATACATGATA CTTGGGTAGC TCAACATTCT TAATAAACAG CCTTTGTATT 1234
 CAGAAATATA AATTGAAATA GATATATATA AAGTTAAAAA AAAAAAAAAA AAA 1287

Fig. 1

	10v	20v	30v	40v	50v
HLASP	MTSCHIAEEHIQKVAIFGGTHGNETLGVFLVKHWLENGAEIQRTGLEVKPF				
	MTSCH:AE:.I:KVAIFGGTHGNETLGVFLVKHWLEN::EIQRTGLEVKPF				
BASPCDNA	MTSCHVAEDPIKKVAIFGGTHGNETLGVFLVKHWLENSTEIQRTGLEVKPF				
	10^	20^	30^	40^	50^
	60v	70v	80v	90v	100v
HLASP	ITNPRAVKKCTRYIDCDLNRIFDLENLGKKK\$EDLPYEVRRAQEINHFLFGP				
	ITNPRAVKKCTRYIDCDLNR:FD ENLGKK.SEDLPYEVRRAQEINHFLFGP				
BASPCDNA	ITNPRAVKKCTRYIDCDLNRVFDPENLGKKK\$EDLPYEVRRAQEINHFLFGP				
	60^	70^	80^	90^	100^
	110v	120v	130v	140v	150v
HLASP	K\$EDSYDIIIFDLHN*TTSNMGCTLILED\$SRNNFLIQMFHYIKT\$SLAPLPCY				
	K\$EDSYDIIIFDLHN*TTSNMGCTLILED\$SRN:FLIQMFHYIKT\$SLAPLPCY				
BASPCDNA	K\$EDSYDIIIFDLHN*TTSNMGCTLILED\$SRNDFLIQMFHYIKT\$SLAPLPCY				
	110^	120^	130^	140^	150^
	160v	170v	180v	190v	200v
HLASP	VYLIEHPSLKYATTRSIKYPVGIEVGPQPGVLRADILDQMRKMIKHALD				
	VYLIEHPSLKYATTRSIKYPVGIEVGPQPGVLRADILDQMRKMI:HALD				
BASPCDNA	VYLIEHPSLKYATTRSIKYPVGIEVGPQPGVLRADILDQMRKMIQHALD				
	160^	170^	180^	190^	200^
	210v	220v	230v	240v	250v
HLASP	FIHHFNEGKEFPPCAIEVYKIIKVDYPRDENGIEIAAIIHPNLQDQDWKPL				
	FIH:FNEGKEFPPCAIEVYKI: KVDYPR:E:GEI:AIHP:LQDQDWKPL				
BASPCDNA	FIHNFNEGKEFPPCAIEVYKIMRKVDYPRNESGEISAIHPKLQDQDWKPL				
	210^	220^	230^	240^	250^
	260v	270v	280v	290v	300v
HLASP	HPGDPMFLTLTDGKTIPLGGDCTVYPVVFVNEAAYYEKKEAFKTTKLTNLAK				
	HP.DP:FLTLTDGKTIPLGGD TVYPVVFVNEAAYYEKKEAFKTTKLTNLAK:				
BASPCDNA	HPEDPVFLTLTDGKTIPLGGDQTVYPVVFVNEAAYYEKKEAFKTTKLTNLAK				
	260^	270^	280^	290^	300^
	310v				
HLASP	SIRCCLH				
	SIR..LH				
BASPCDNA	SIRSSLH				
	310^				

Fig. 2

09965807-100101

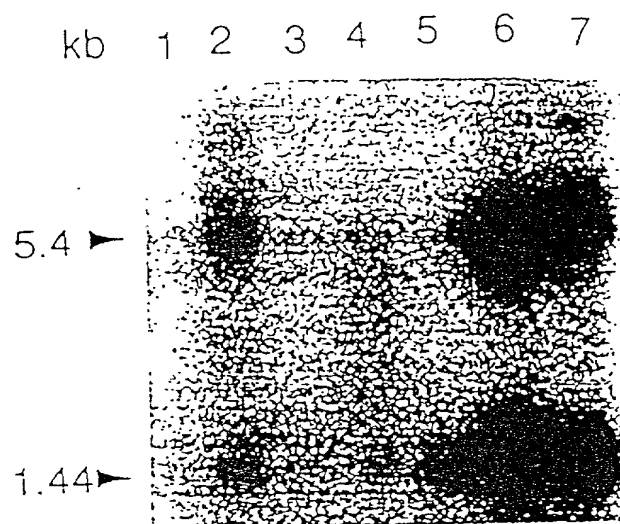


Fig. 3

Kaul et al.

FOOT-20859660

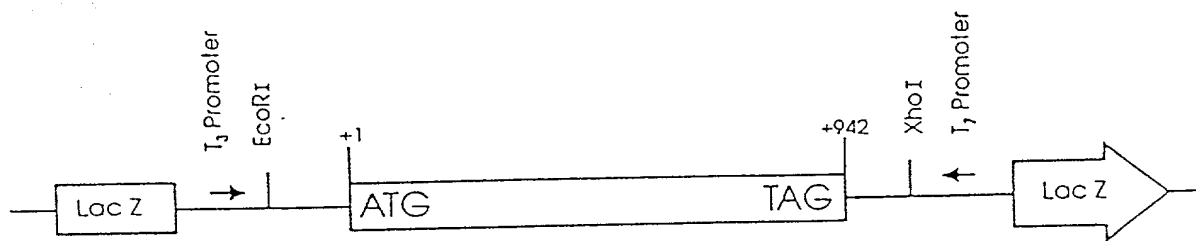


Fig. 4

Kaul et al.

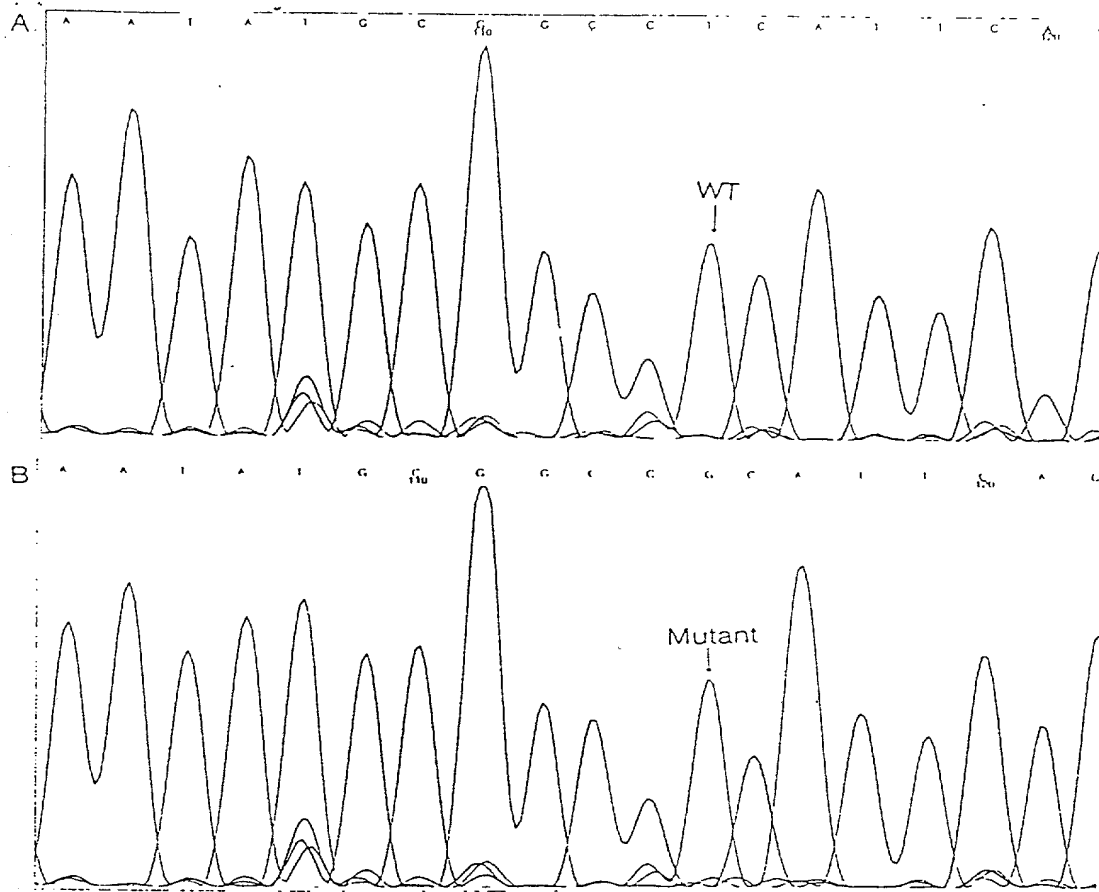


Fig. 5

Kaul et al.

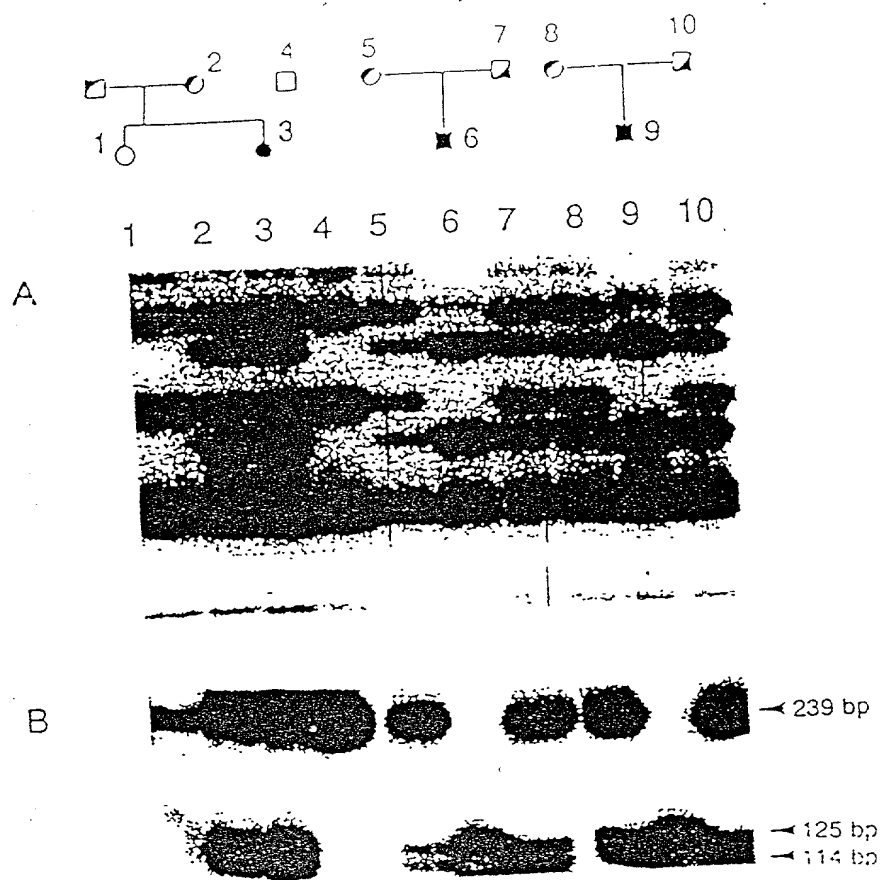


Fig. 6
Kaul et al.

8/26/93

MAPSEQ V5.33 .HASP.SEQ(1,1277) Reading frames: 1 Enzyme file ALL.ENZ LinPage 1

EAM	M	E	M	N	DSNDSBB
ASA	B	C	N	L	STCSESS
MPE	O	5	L	A	AYOACAA
1E3	2	7	1	4	11111JJ

ATGACTTCTTGTCACATTGCTGAAGAACATATACAAAAGGTTGCTATCTTTGGAGGAACC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60

TACTGAAGAACAGTGTAACGACTTCTTGTATATGTTTTCCAACGATAGAAACCTCCTTGG
 m t s c h i a e e h i q k v a i f g g t

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

N	A	BBH	TSM	RM	H	HHD	TH
L	L	SCP	RPS	MA	I	HAD	FN
A	U	AAA	UOE	AE	N	AEE	IF
3	1	W72	911	11	P	121	11

CATGGGAATGAGCTAACCGGAGTATTTCTGGTTAAGCATTGGCTAGAGAATGGCGCTGAG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 120

GTACCCTTACTCGATTGGCCTCATAAAGACCAATTCGTAACCGATCTCTTACCGCGACTC
 h g n e l t g v f l v k h w l e n g a e

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

M	B
N	P
L	M
1	1

ATTCAGAGAACAGGGCTGGAGGTAAAACCATTTATTACTAACCCCAGAGCAGTGAAGAAG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180

TAAGTCTCTTGTCCCGACCTCCATTTTGGTAAATAATGATTGGGGTCTCGTCACTTCTTC
 i q r t g l e v k p f i t n p r a v k k

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

Start site

00965807.100101

Restriction Sites

Fig. 7(a)

1-67 15
 68-131 16
 132-141 17

CR	M	M	TH
SS	B	A	FN
PA	O	E	IF
61	2	3	11

TGTACCAGATATATTGACTGTGACCTGAATCGCATTTTTGACCTTGAAAATCTTGGCAAA
-----+-----+-----+-----+-----+-----+-----+
ACATGGTCTATATAACTGACACTGGACTTAGCGTAAAAACTGGAACCTTTTAGAACCGTTT 240
c t r y i d c d l n r i f d l e n l g k

-----+-----+-----+-----+-----+-----+-----+
NM BN
DB AS
EO NP
12 22
/
AAAATGTCAGAAGATTTGCCATATGAAGTGAGAAGGGCTCAAGAAATAAATCATTATTT
-----+-----+-----+-----+-----+-----+-----+ 300
TTTTACAGTCTTCTAAACGGTATACTTCACTCTTCCCGAGTTCTTTATTTAGTAAATAAA
k m s e d l p y e v r r a q e i n h l f

-----+-----+-----+-----+-----+-----+-----+
A TH M S
V FN B P
A IF O O
2 11 2 1
/
GGTCCAAAAGACAGTGAAGATTCCTATGACATTATTTTTGACCTTCACAACACCACCTCT
-----+-----+-----+-----+-----+-----+-----+ 360
CCAGGTTTTCTGTCACTTCTAAGGATACTGTAATAAAAACTGGAAGTGTTGTGGTGGAGA
g p k d s e d s y d i i f d l h n t t s

Fig. 7(b)

09965807-100101

MN A	HBN M	TH E AS	TM
NL P	GSS N	FN C PC	RS
LA L	IIP L	IF R YR	UE
13 1	AH2 1	11 2 11	91

AACATGGGGTGCACCTCTTATTCTTGAGGATTCAGGAATAACTTTTAAATTCAGATGTTT
 -----+-----+-----+-----+-----+-----+-----+-----+
 TTGTACCCACGTGAGAATAAGAACTCCTAAGGTCCTTATTGAAAAATTAAGTCTACAAA

420

n m g c t l i l e d s r n n f l i q m f

TM	N	M F	E
RS	L	A O	C
UE	A	E K	O
91	4	2 1	B

CATTACATTAAGACTTCTCTGGCTCCACTACCCTGCTACGTTTATCTGATTGAGCATCCT
 -----+-----+-----+-----+-----+-----+-----+-----+
 GTAATGTAATTCTGAAGAGACCGAGGTGATGGGACGATGCAAATAGACTAACTCGTAGGA

480

h y i k t s l a p l p c y v y l i e h p

S	M	A
F	N	V
A	L	A
N	1	2

TCCCTCAAATATGCGACCACTCGTTCCATAGCCAAGTATCCTGTGGGTATAGAAGTTGGT
 -----+-----+-----+-----+-----+-----+-----+-----+
 AGGGAGTTTATACGCTGGTGAGCAAGGTATCGGTTTCATAGGACACCCATATCTTCAACCA

540

s l k y a t t r s i a k y p v g i e v g

Fig.7(c)

09965807-100101

D	M	M D	A	E	BMDD	TM
D	N	N D	L	C	IBPP	RS
E	L	L E	U	R	NONN	UE
1	1	1 1	1	V	1121	91

115 bp

CCTCAGCCTCAAGGGGTTCTGAGAGCTGATATCTTGGATCAAATGAGAAAAATGATTAAA
-----+-----+-----+-----+-----+-----+-----+
GGAGTCGGAGTTCCCCAAGACTCTCGACTATAGAACCTAGTTTACTCTTTTTACTAATTT

600

p q p q g v l r a d i l d q m r k m i k

-----+-----+-----+-----+-----+-----+-----+

NN	HMHM
SL	INHN
PA	NLAL
H3	P111

/ CATGCTCTTGATTTTATACATCATTTCAATGAAGGAAAAGAATTTCCCTCCCTGCGCCATT
-----+-----+-----+-----+-----+-----+-----+
GTACGAGAACTAAAATATGTAGTAAAGTTACTTCCTTTTCTTAAAGGAGGGACGCGGTAA

660

h a l d f i h h f n e g k e f p p c a i

-----+-----+-----+-----+-----+-----+-----+

E	BSBNXSASSBBHNSB	FF	F	IF
C	SESCMMVCESSPCCB	OO	O	TN
P	ACAIAAARCAAIRV	KK	K	AU
1	J1J111111JJ2111	11	1	1H

546p

GAGGTCTATAAAATTATAGAGAAAGTTGATTACCCCCGGGATGAAAATGGAGAAATTGCT
-----+-----+-----+-----+-----+-----+-----+
CTCCAGATATTTTAATATCTCTTTCAACTAATGGGGGCCCTACTTTTACCTCTTTAACGA

720

e v y k i i e k v d y p r d e n g e i a

-----+-----+-----+-----+-----+-----+-----+

C 693 > a
Y 231 > X

Fig. 7 (d)

00655807-100101

S	PBMDD	F	ESASBBSBXBNMDDDB
F	SIBPP	O	CEPCSSFIHALBPPI
C	TNONN	K	RCYRAAANOMAONNN
1	11121	1	2111JJN12141211

//////

GCTATCATCCATCCTAATCTGCAGGATCAAGACTGGAAACCACTGCATCCTGGGGATCCC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 780
 CGATAGTAGGTTAGGATTAGACGTCCTAGTTCTGACCTTTGGTGACGTAGGACCCCTAGGG

a i i h p n l q d q d w k p l h p g d p

-----+-----+-----+-----+-----+-----+-----+-----+-----+

(90 bp) ✓

N	TM	B	MDBBBDBMA	BBAB	CR	CR
L	RS	B	BPBSPPIBL	SSCS	SS	SS
A	UE	V	ONSCUNNOW	ILIM	PA	PA
3	91	2	121911122	Y112	61	61

/ // // //

ATGTTTTTAACTCTTGATGGGAAGACGATCCCACTGGGCGGAGACTGTACCGTGTACCCC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 840
 TACAAAAATTGAGAACTACCCTTCTGCTAGGGTGACCGCCTCTGACATGGCACATGGGG

m f l t l d g k t i p l g g d c t v y p

-----+-----+-----+-----+-----+-----+-----+-----+-----+

SM	HIFA	H A
PN	ATNC	I L
OL	EAUI	N U
11	31H1	3 1

/ * //

GTGTTTGTGAATGAGGCCGCATATTACGAAAAGAAAGAAGCTTTTGCAAAGACAACCTAAA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
 CACAAACACTTACTCCGGCGTATAATGCTTTTCTTTCTTCGAAAACGTTTCTGTTGATTT

v f v n e a a y y e k k e a f a k t t k

-----+-----+-----+-----+-----+-----+-----+-----+-----+

0354>C
E285>A

Fig. 7(e)

S	B	HIF	E	A
P	B.	NTN	C	L
O	V	FAU	1	U
1	1	31H	5	1

Ala 305>L
C9147A

CTAACGCTCAATGCAAAAAGTATTTCGCTGCTGTTTACATTAGAAATCACTTCCAGCTTAC
-----+-----+-----+-----+-----+-----+-----+-----+
GATTGCGAGTTACGTTTTTCATAAGCGACGACAAATGTAATCTTTAGTGAAGGTCGAATG

960

l t l n a k s i r c c l h . k s l p a y

-----+-----+-----+-----+-----+-----+-----+-----+

RM	A	ATM
MA	L	FRS
AE	U	LUE
11	1	291

ATCTTACACGGTGTCTTACAAATTCTGCTAGTCTGTAAGCTCCTTAAGAGTAGGGTTGTG
-----+-----+-----+-----+-----+-----+-----+-----+
TAGAATGTGCCACAGAATGTTTAAGACGATCAGACATTTCGAGGAATTCTCATCCCAACAC

1020

i l h g v l q i l l v c k l l k s r v v

-----+-----+-----+-----+-----+-----+-----+-----+

B	A	RM	H	D	S
S	L	MA	N	D	F
P	U	AE	F	E	A
W	1	11	3	1	N

CCTTATTCAACTGCATACATAGCTCCTAGCACAGTGCCTTATTCGGTAGGCATCTAAGCA
-----+-----+-----+-----+-----+-----+-----+-----+
GGAATAAGTTGACGTATGTATCGAGGATCGTGTCACGGAATAAGCCATCCGTAGATTCGT

1080

p y s t a y i a p s t v p y s v g i . a

-----+-----+-----+-----+-----+-----+-----+-----+

Fig. 7(f)

0095807-100101

TM	ATM	PATM	TDM	E	A
RS	SRS	ASRS	RRS	C	L
UE	EUE	CEUE	UAE	R	U
91	191	1191	911	V	1

/ // /// /
AATTTCTTAAATTAATTAATATATCTTTAAAGATATCATATTTTATGTATGTAGCTTATT
-----+-----+-----+-----+-----+-----+-----+ 1140
TTAAAGAATTTAATTAATTATATAGAAATTTCTATAGTATAAAATACATACATCGAATAA

n f l n . l i y l . r y h i l c m . l i

X	N	A
M	L	L
N	A	U
1	3	1

CAAAGAAGTGTTTCCTATTTCTATATAGTTTATTATACATGATACTTGGGTAGCTCAACA
-----+-----+-----+-----+-----+-----+-----+ 1200
GTTTCTTCACAAAGGATAAAGATATATCAAATAATATGTACTATGAACCCATCGAGTTGT

q r s v s y f y i v y y t . y l g s s t

✓ TM	✓ TM
RS	RS
UE	UE
91	91
/	/

TTCTTAATAAACAGCCTTTGTATTCAGAATATAAAATTGAAATAGATATATATAAAGTTA
-----+-----+-----+-----+-----+-----+-----+ 1260
AAGAATTATTTGTCGAAACATAAGTCTTATATTTTAACTTTATCTATATATATTTCAAT

f l i n s l c i q n i k l k . i y i k l

AAAAAAAAAAAAAAAAAAAA
-----+-----+-----+ 1277
TTTTTTTTTTTTTTTTTTT

k k k k k k

Fig. 7(g)

S	B	HIF	E	A
P	B	NTN	C	L
O	V	FAU	1	U
1	1	31H	5	1

Ala 305
C91-47

CTAACGCTCAATGCAAAAAGTATTCGCTGCTGTTTACATTAGAAATCACTTCCAGCTTAC
 -----+-----+-----+-----+-----+-----+-----+-----+
 GATTGCGAGTTACGTTTTTCATAAGCGACGACAAATGTAATCTTTAGTGAAGGTGGAATG

960

l t l n a k s i r c c l h .

-----+-----+-----+-----+-----+-----+-----+-----+

RM	A	ATM
MA	L	FRS
AE	U	LUE
11	1	291

ATCTTACACGGTGTCTTACAAATTCTGCTAGTCTGTAAGCTCCTTAAGAGTAGGGTTGTG
 -----+-----+-----+-----+-----+-----+-----+-----+
 TAGAATGTGCCACAGAATGTTTAAGACGATCAGACATTGAGGAATTCTCATCCCAACAC

1020

-----+-----+-----+-----+-----+-----+-----+-----+

B	A	RM	H	D	S
S	L	MA	N	D	F
P	U	AE	F	E	A
W	1	11	3	1	N

CCTTATTCAACTGCATACATAGCTCCTAGCACAGTGCCTTATTCGGTAGGCATCTAAGCA
 -----+-----+-----+-----+-----+-----+-----+-----+
 GGAATAAGTTGACGTATGTATCGAGGATCGTGTACGGAATAAGCCATCCGTAGATTTCGT

1080

-----+-----+-----+-----+-----+-----+-----+-----+

0995807 100101

Fig. 7(f)

SEQ V5.33 HASP.SEQ(1,1277) Reading frames: 1 Enzyme file ALL.ENZ LinPag

TM	ATM	PATM	TDM	E	A
RS	SRS	ASRS	RRS	C	L
UE	EUE	CEUE	UAE	R	U
91	191	1191	911	V	1
/	//	///	/		

AATTTCTTAAATTAATTAATATATCTTTAAAGATATCATATTTTATGTATGTAGCTTATT
-----+-----+-----+-----+-----+-----+-----+
TTAAGAATTTAATTAATTATATAGAAATTTCTATAGTATAAAATACATACATCGAATAA

1140

-----+-----+-----+-----+-----+-----+-----+

X	N	A
M	L	L
N	A	U
1	3	1

CAAAGAAGTGTTTCCTATTTCTATATAGTTTATTATACATGATACTTGGGTAGCTCAACA
-----+-----+-----+-----+-----+-----+-----+
GTTTCTTCAGAAAGGATAAAGATATATCAAATAATATGTACTATGAACCCATCGAGTTGT

1200

-----+-----+-----+-----+-----+-----+-----+

TM	TM
RS	RS
UE	UE
91	91
/	/

TTCTTAATAAACAGCCTTTGTATTTCAGAATATAAAATTGAAATAGATATATATAAAGTTA
-----+-----+-----+-----+-----+-----+-----+
AAGAATTATTTGTCGAAACATAAGTCTTATATTTTAACTTTATCTATATATATTTCAAT

1260

AAAAAAAAAAAAAAAAAAAA

-----+-----+-----+-----+-----+-----+-----+
TTTTTTTTTTTTTTTTTTTT 1277

-----+-----+-----+-----+-----+-----+-----+

Fig. 7(g)